

Genotypes: a concept to help integrate molecular phylogenetics and taxonomy

PROSANTA CHAKRABARTY

Museum of Natural Science, Louisiana State University, Baton Rouge, LA 70803, USA. E-mail: prosanta@lsu.edu

In order to better integrate molecular phylogenetics and taxonomy, genetic sequences from type materials should be explicitly identified in publications using a consistent nomenclature. Despite relatively frequent sampling of sequences from types (particularly topotypes—samples from the type locality), the practice of explicitly noting that these materials were sampled is uncommon. Because of the lack of an explicit nomenclature tied to taxonomy, the existence of genetic “type sequences” is obscured. Also hindering progress in taxonomy is the increasingly uncommon practice of reporting locality and voucher information (e.g., GPS coordinates, museum catalog numbers) on repositories such as GenBank. To remedy this problem and bring awareness to the situation, I propose the use of the term “genotype” as a label for any sequence data from types (including from holotypes, secondary types, topotypes, etc.).

For nomenclatural purposes, genetic sequences from a holotype should be referred to as a “hologenotype” (from: **holotype** and **genotype**), sequences from a topotype will be a “topogenotype,” and so forth. In addition, the genetic marker(s) used should also be incorporated into the nomenclature (e.g., paragenotype ND2). Having the term “genotype” linked to the traditional taxonomic nomenclature will greatly simplify the process of finding comparative data. If the term “genotypes” is included in the text of a publication, academic search engines such as Google Scholar, Scopus and Web of Science will be capable of finding genetic sequences from types much more rapidly than is currently possible. These genotypes can then be used immediately in comparisons and compiled into databases such as GenBank, GBIF (Global Biodiversity Information Facility), or taxon specific sites such as the *World Spider Catalog* (Platnick 2010), *Amphibian Species of the World* (Frost 2010) or the *Catalog of Fishes* (Eschmeyer 2010). Having a database of genotypes will create a new dimension by which taxa can be compared in decisions related to the taxonomic status of species.

The influence of molecular data on taxonomy is poorly understood as evidenced by the growing rift between molecular biologists and traditional morphologists (Mooi and Gill 2010; Chakrabarty 2010a; Smith in press; Wiley *et al.* in press). However, the divide between these two schools may be repaired by the proposed approach of identifying and utilizing genotypes. Enacting this approach will greatly simplify many aspects of modern taxonomy. By making gene sequences from type materials readily available, researchers will be able to rapidly compare what they suspect to be undescribed species, new populations, or species in synonymy.

Type materials remain essential for taxonomic comparisons, but sequence data have not been fully incorporated into these comparisons, or into the tradition of taxonomy. Access to sequence data from type materials will allow for an objective basis for taxonomic comparisons. Judgments about the amount of sequence divergence that is required to differentiate between taxa remains with the individual researchers and their preferred species concepts.

The genotypes approach is not a call to take tissue samples from newly collected primary types (particularly the holotype, which should be maintained as the morphological ideal). Although primary or secondary types are sometimes sequenced (Sparks and Dunlap 2004; Meyer *et al.* 2008; Chakrabarty 2010b), the lack of an identifying keyword or flag for tracking these sequences makes them exceedingly difficult to discover. Researchers should also consider using specimens that are DNA vouchers as part of a paratype series in the description of a new species. Having a paratype that is also a DNA voucher will ensure that genetic materials are available from the type series. Having paragenotypes will allow comparisons of type materials that are not restricted to morphology. Currently the vast majority of taxonomic decisions (e.g., synonyms, descriptions) rely solely on morphology. Morphological comparisons should always be part of this process; however, adding a molecular component will add a new dimension to taxonomic research.

Most genotypes will likely come from topotypes (because these can be sampled after the original description is published), and it is these topogenotypes that will provide most of the primary comparative data to molecular taxonomists. Availability of genotypes to researchers will solve multiple taxonomic problems by providing a new layer of objectivity to taxonomy. Disputes over the validity of a given taxon can be rapidly solved with comparisons using genotypes. For instance, if sequence data from a holotype are available, material thought to be distinct can be quickly compared. Similarly, topogenotypes can be helpful to researchers in cases where primary or secondary types are unavailable for sequencing. It will also become quite simple to compare widely separated populations from taxa with extensive or disjunct distributions.

Researchers should not restrict their choice of representative genotypes to typical barcoding genes. Although researchers are free to choose anything from a portion of a gene to an entire genome as a genotype, they should be mindful that the level of comprehensiveness of genetic comparisons is directly proportional to the amount of data made available. Just as there is variation among morphological type specimens, there will be sequence variation in representative genotypes of a given taxon and researchers should take this variation into account when considering taxonomic changes. The gold standard of the genotype will be the hologenotype that is represented by that holotype's entire genome. However, it is likely that initially most genotypes will be topogenotypes represented by a relatively small percentage of the topotype's genome. Ultimately, the genotypes approach will promote all forms of taxonomic research as molecular phylogenetics becomes integrated with taxonomy and as technology improves in molecular biology.

Systematists who use molecular tools today sometimes find putative novel taxa based on molecular phylogenies that reveal previously undescribed or cryptic diversity. These data are sometimes used to justify the descriptions of a new taxon even though the sequence data itself is rarely used in the actual description. Some taxonomic journals do not even require the submission of genetic sequences to GenBank (e.g., *Zootaxa*). Unfortunately, sequences from types are rarely made available or are often obscured within the appendices of publications. A database of genotypes will provide a source for these comparative data and promote the field of molecular taxonomy. It will also provide an accessible inventory to help all taxonomists answer difficult questions with molecular tools. The first step to reaching this next level of taxonomy is the straightforward adaptation of this proposed approach and genotype nomenclature.

Acknowledgments

I would like to thank Dr. Larry Page and Dr. Bill Eschmeyer for their comments on this manuscript and for several discussions on the subject. I would also like to thank the students in the Systematics Discussion Group at LSU for useful comments. This work was supported by startup funds from Louisiana State University and the Museum of Natural Science as well as the National Science Foundation DEB 0916695.

Literature cited

- Chakrabarty, P. (2010a) The transitioning state of systematic ichthyology. *Copeia*, 2010, 513–515.
- Chakrabarty, P. (2010b) Status and phylogeny of Milyeringidae (Teleostei:Gobiiformes), with the description of a new blind cave-fish from Australia, *Milyeringa brooksi*, new sp. *Zootaxa*, 2557, 19–28.
- Eschmeyer, W.N. (ed.) (2010) *Catalog of Fishes, electronic version* California Academy of Science, CA. Available from: <http://research.calacademy.org/ichthyology/catalog/fishcatmain.asp> (26 August 2010).
- Frost, D.R. (2010) *Amphibian Species of the World: an Online Reference. Version 5.4*, American Museum of Natural History, NY. Available from: <http://research.amnh.org/vz/herpetology/amphibia/> (8 April, 2010).
- Meyer, M.J., Price, R.D. & Johnson, K.P. (2008) A new species of *Picicola* Clay and Meinertzhagen, 1938 (Phthiraptera: Ischnocera) parasitic on the Rufous-sided Broadbill (Passeriformes: Eurylaimidae) in Ghana. *Zootaxa*, 1762, 63–68.
- Mooi, R.D. & Gill, A.C. (2010) Phylogenetics without synapomorphies – A crisis in fish systematics: time to show some character. *Zootaxa*, 2450, 26–40.
- Platnick, N.I. (2010) *The World Spider Catalog, version 11.0*, American Museum of Natural History, NY. Available from: <http://research.amnh.org/entomology/spiders/catalog/index.html> (7 September, 2010).
- Smith, W.L. (in press) Promoting resolution of the percomorph bush: a reply to Mooi and Gill. *Copeia*.
- Sparks, J.S. & Dunlap, P.V. (2004) A clade of non-sexually dimorphic ponyfishes (Teleostei: Perciformes: Leiognathidae): phylogeny, taxonomy, and description of a new species. *American Museum Novitates*, 3459, 1–21.
- Wiley, E.O., Chakrabarty, P., Craig, M.T., Davis, M.P., Holcroft, N.I., Mayden, R.L. & Smith, W.L. (in press) Will the real phylogeneticists please stand up? *Zootaxa*.